



OPEN ACCESS

EDITED BY

Ping Zhou,
University of Health and Rehabilitation
Sciences, China

REVIEWED BY

Helge Arne Wurdemann,
University College London, United
Kingdom
Sen Qiu,
Dalian University of Technology, China
Natalia Sanchez,
Chapman University, United States

*CORRESPONDENCE

Jiaen Wu
wujiae@ethz.ch

SPECIALTY SECTION

This article was submitted to
Neural Technology,
a section of the journal
Frontiers in Neuroscience

RECEIVED 24 June 2022

ACCEPTED 18 November 2022

PUBLISHED 08 December 2022

CITATION

Wu J, Maurenbrecher H, Schaer A,
Becsek B, Awai Easthope C,
Chatzipirpiridis G, Ergeneman O,
Pané S and Nelson BJ (2022) Human
gait-labeling uncertainty and a hybrid
model for gait segmentation.
Front. Neurosci. 16:976594.
doi: 10.3389/fnins.2022.976594

COPYRIGHT

© 2022 Wu, Maurenbrecher, Schaer,
Becsek, Awai Easthope,
Chatzipirpiridis, Ergeneman, Pané and
Nelson. This is an open-access article
distributed under the terms of the
[Creative Commons Attribution License
\(CC BY\)](https://creativecommons.org/licenses/by/4.0/). The use, distribution or
reproduction in other forums is
permitted, provided the original
author(s) and the copyright owner(s)
are credited and that the original
publication in this journal is cited, in
accordance with accepted academic
practice. No use, distribution or
reproduction is permitted which does
not comply with these terms.

Human gait-labeling uncertainty and a hybrid model for gait segmentation

Jiaen Wu^{1,2*}, Henrik Maurenbrecher², Alessandro Schaer²,
Barna Becsek², Chris Awai Easthope³, George Chatzipirpiridis²,
Olgac Ergeneman², Salvador Pané¹ and Bradley J. Nelson¹

¹Multi-Scale Robotics Lab, ETH Zurich, Zurich, Switzerland, ²Magnes AG, Zurich, Switzerland,
³Cereneo Foundation, Center for Interdisciplinary Research (CEFIR), Vitznau, Switzerland

Motion capture systems are widely accepted as ground-truth for gait analysis and are used for the validation of other gait analysis systems. To date, their reliability and limitations in manual labeling of gait events have not been studied.

Objectives: Evaluate manual labeling uncertainty and introduce a hybrid stride detection and gait-event estimation model for autonomous, long-term, and remote monitoring.

Methods: Estimate inter-labeler inconsistencies by computing the limits-of-agreement. Develop a hybrid model based on dynamic time warping and convolutional neural network to identify valid strides and eliminate non-stride data in inertial (walking) data collected by a wearable device. Finally, detect gait events within a valid stride region.

Results: The limits of inter-labeler agreement for key gait events heel off, toe off, heel strike, and flat foot are 72, 16, 24, and 80 ms, respectively; The hybrid model's classification accuracy for stride and non-stride are 95.16 and 84.48%, respectively; The mean absolute error for detected heel off, toe off, heel strike, and flat foot are 24, 5, 9, and 13 ms, respectively, when compared to the average human labels.

Conclusions: The results show the inherent labeling uncertainty and the limits of human gait labeling of motion capture data; The proposed hybrid-model's performance is comparable to that of human labelers, and it is a valid model to reliably detect strides and estimate the gait events in human gait data.

Significance: This work establishes the foundation for fully automated human gait analysis systems with performances comparable to human-labelers.

KEYWORDS

gait labeling uncertainty, limit of agreement, convolutional neural network, dynamic time warping, automatic gait segmentation, wearable inertial sensors, gait event detection, human activity recognition (HAR)

1. Introduction

Human walking is a complex dynamic process. It integrates multiple interdependent sensory inputs and requires coordination of the motor outputs to achieve efficient, stable, and adaptive locomotion. The deterioration of human gait can severely affect people's mobility, reduce a person's level of activity, and affect their quality of life. In general, gait can be described by its temporal and spatial characteristics such as stride time, stride length, stride velocity, swing time, stance time, and their corresponding variability as well as left and right symmetry metrics (Del Din et al., 2016). Probing and evaluating these gait characteristics is important not only in the field of rehabilitation (Richards et al., 1999; Eng and Tang, 2007; Selves et al., 2020), but also in the early detection of neurological diseases, such as Parkinson's disease (Lees, 1992; Brognara et al., 2019), Alzheimer's disease (Gras et al., 2015; Mc Ardle et al., 2019), and tremor syndromes (Stolze et al., 2001; Hoskovicová et al., 2013). Some specific gait patterns even appear in the earliest stage of neurological diseases before other signs and symptoms are evident, providing significant clues for early diagnosis and treatment (Jacobi et al., 2012; Nürnberger et al., 2015; Pistacchi et al., 2017). Therefore, quantitative gait characteristics can be used as a digital biomarker for monitoring and assessing people's health status.

To assess and evaluate an individual's gait by temporal and spatial characteristics quantitatively, gait is segmented into repeating cycles. Each cycle starts when the foot leaves the ground commencing the swing phase and is followed by the stance phase, where the foot is on the ground again, which lasts until the next gait cycle starts (DeLisa, 1998). These two main phases can be further divided into eight sub-phases by gait events within one gait cycle: initial contact, loading response, mid-stance, terminal stance, pre-swing, initial swing, mid-swing, and terminal swing. Key gait events segmenting these phases are heel strike (HS) and toe off (TO). Additional events of interest include heel off (HO) and flat foot (FF), which give insights into the transition periods between the two main gait phases, swing and stance. For pathological gait, the preceding sequence and the time allocation of gait phases can vary compared to healthy people, and some gait events or phases might even be missing completely (Senanayake and Senanayake, 2010; Meng et al., 2013). Observing these variations in gait events long-term, during daily activities provides a powerful tool for clinical gait assessment, evaluation, and diagnosis (Atrsaei et al., 2021), such as identifying early Parkinson's disease (PD) (Rehman et al., 2019), capturing freezing of gait (Palmerini et al., 2017; Mancini et al., 2019), monitoring PD symptoms (Heijmans et al., 2019), and the Levodopa response (Pulliam et al., 2017). In light of the ongoing shift of rehabilitation and early diagnosis from physical visits of care facilities toward remote monitoring systems, the understanding of real-world gait data

gains significant importance (Warmerdam et al., 2020). Recent studies have investigated the relationships between clinical gait assessments and home-based monitoring systems, and have described an additive component for real-world data (Shah et al., 2020; Atrsaei et al., 2021). A reliable method for continuous remote monitoring of gait quantity and quality affords new perspectives of assessing the real-world impact of interventions and of investigating the transfer of rehabilitation progress to out-patient behavior.

To capture the quality of gait, a significant research effort has been made toward achieving a comprehensive understanding of spatio-temporal gait characteristics, with the aid of various systems such as motion capture systems (Mihradi et al., 2011; Carse et al., 2013), pressure mats (Yu et al., 2010; Papavasileiou et al., 2017), and wearable sensors (Tunca et al., 2017; Zhang et al., 2019; Wu et al., 2021; Celik et al., 2022). Compared to traditional subjective visual observations by therapists, instrumented assessments allow for an objective assessment of gait characteristics (Laughman et al., 1984). Motion capture systems give a detailed evaluation of motion but require a complex stationary set-up involving cameras and the collected data needs to be post-processed (manually) to identify relevant gait characteristics. This makes those gait systems very expensive to operate. Moreover, motion capture systems cannot track people outside the laboratory and therefore the time and place of the data collection are severely constrained as the patient must physically visit a gait laboratory. An alternative is pressure mats, which measure the interface pressure when people are treading on them. While the setup has reduced complexity, gait measurement is still confined to a gait laboratory and the available range of gait parameters and information from pressure mats is limited. Both motion capture systems and pressure mats offer a constrained walking environment, allowing individuals to focus their attention solely on ambulation. This may lead to biased results, as the recorded gait is not representative of the patient's typical gait. Additionally, the measurements often take place under the supervision of a health professional, which may further skew the collected data (Kaye et al., 2012; Robles-García et al., 2015). These costly and laboratory-limited data acquisition systems have led to a growing urge for cheaper and mobile systems, i.e., wearable gait analysis system (WGAS) (Tunca et al., 2017; Zhang et al., 2019; Wu et al., 2021). WGAS are not only cost-effective and portable, but can also be used for a longer period of time of continuous monitoring, allowing for unbiased measurements during daily routines.

For validation of WGAS's algorithms, optical motion capture systems are the most commonly used reference systems and are considered as the gold standard (Caldas et al., 2017; Kobsar et al., 2020). They employ either retro-reflective (passive) or infrared emitting (active) markers attached to different locations on the lower limbs. These markers are tracked through multiple

cameras to record their spatial trajectories. The key gait events are then identified by technicians who visually examine the markers' spatial trajectories or by automated event detection algorithms based on markers' position and velocity (Kidziński et al., 2019; Lempereur et al., 2020) for further gait parameter estimation. Although the aforementioned automated detection algorithms are time efficient, they are not suited for the patient population due to the high variability that is intrinsic to pathological gait. Thus, the manual labeling of motion capture system data is still the most widely accepted benchmark for validation of WGAS algorithms for gait analysis (Caldas et al., 2017; Tunca et al., 2017; Hsu et al., 2018). Nonetheless, the uncertainty of manual labels of motion capture system data and the limit of detection have never been considered and evaluated. The uncertainty in manual labels mainly comes from two sources: 1) the lack of consistency within a labeler i.e., intra-labeler inconsistency; and 2) the lack of consistency among different labelers i.e., inter-labeler inconsistency. The label quality of gait events and acquisition results are highly dependent on the labeler's personal perspective and experience level. The uncertainty of labels in gait data resulting from intra-labeler and inter-labeler inconsistency can have a direct impact on the reliability of gait event labels, and further affect the accuracy of gait parameter estimation and the decision by doctors. Therefore, one objective of this study is to investigate the uncertainty in manual labeling of gait events in the well-accepted benchmark, i.e., optical motion capture system data.

To estimate spatial and temporal gait parameters from WGAS, different algorithms have been developed. The first group of gait analysis methods relies on straight gait event detection in long-term sensor signals (Agostini et al., 2013). Gait events can be identified by different techniques, including rule-based (Zhu et al., 2012; Zhao et al., 2017) and machine learning methods (Zhen et al., 2019; Liu et al., 2021). The second group of methods splits gait analysis into two stages: first segmenting the sensor data into strides, and then searching for gait events within each found stride. The advantages of identifying single strides first are that it is more robust for pathological gait where some gait events may be missing (Agostini et al., 2013), and more efficient for long-term gait recording processing since it rejects non-stride data in the first stage, reducing the computational cost (Ullrich et al., 2020). This two-stage gait analysis has been investigated in literature by many techniques, such as threshold/peak detection methods (Hickey et al., 2016; Ullrich et al., 2020), template matching methods (Barth et al., 2015; Oudre et al., 2018) and machine learning-based methods (Martindale et al., 2021). However, robust algorithms for automatic detection of stride and rejection of non-stride data from long-term free-living walking data are still needed for effective gait analysis. Therefore, in this paper, a hybrid approach based on dynamic time warping (DTW) and convolutional neural network (CNN) is proposed to automatically identify stride

signals and eliminate non-stride signals from long-term inertial walking data.

The contribution of this study is two-fold: 1) The uncertainty in terms of the limit of agreement (LOA) in manual labeling of human walking data from an optical motion capture system is assessed and analyzed. Those manual labels are further used as benchmarks for the validation of the proposed gait analysis algorithm. Their inherent variances serve as a baseline for the validation study. 2) A hybrid approach based on DTW and CNN named *stepperNet* is developed for gait cycle identification in long-term walking. A gait event detection algorithm is later employed to segment each stride data-slice into sub-phases. The results of gait event detection are compared to the average manual labels obtained from the work done in contribution 1).

2. Materials and methods

2.1. Data acquisition

2.1.1. Participants of the study

Three groups \mathcal{G}_i with $i \in \{1, 2, 3\}$ of healthy participants were recruited in this study. The first group \mathcal{G}_1 consisted of four healthy participants (all males, with an average age of 39 ± 7.07 years), who were instructed to walk on a treadmill. They are labeled as participant $S_1 - S_4$. The dataset of this group was used for the assessment of manual labeling and validation of gait event detection algorithm. The second group \mathcal{G}_2 consisted of (distinct) four healthy participants (all males, with an average age of 31.5 ± 2.06 years), who were instructed to walk for several minutes at various gait regimes. They are labeled as participant $S_5 - S_8$. The dataset of this group was used for *stepperNet* training and validation. The third group \mathcal{G}_3 consisted of another five labelers (four males and one female, with an average age of 29.6 ± 1.67 years). Note that the \mathcal{G}_3 is not a walking group but a group of labelers. To minimize the effect of different labeling experience of labelers and to keep them with the same level of labeling experience, the five labelers were well-trained for the same amount of time to manually label the gait events on the software of the motion capture system. They are labeled as participant $S_9 - S_{13}$.

All participants of \mathcal{G}_1 and \mathcal{G}_2 had no known injuries or abnormalities that affected their gait and were asked to walk as much as possible in the same way they walk in their daily lives. Written informed consent was provided by all participants and this study was conducted in accordance with Good Clinical Practice guidelines and the Declaration of Helsinki after receiving a declaration of clearance from the local ethics committee (BASEC Nr Req-2019-00715).

2.1.2. Wearable gait analysis system

Gait data were recorded with the WGAS as shown in Figure 1. It consists of a pair of shoes with custom-developed embedded electronics. The electronics include a



FIGURE 1
Schematic illustration of WGAS ecosystem embedding sensors in the outsole of the shoes.

wireless-enabled (BLE and Wi-Fi), a dual-core MCU (ESP32, Espressif Systems CO. Ltd., Shanghai, China) running a custom FreeRTOS-based¹ C firmware for handling communication with a custom iOS application, and on-board sensors (LSM6DSM and LSM303AGR, STMicroelectronics, Geneva, Switzerland). The system allows logging sensor data to a local SD card. The WGAS also includes vibration-motors for haptic-feedback.

According to the needs of this research, the range of the accelerometer and gyroscope are set to $\pm 8g$ and $\pm 2,000^\circ/s$, respectively. The sampling frequency of data acquisition is set to 100 Hz. The sensor's body frame, in which the acceleration, angular rate, and magnetic field are measured, can be seen in Figure 1. The x -axis is aligned with the shoe's major axis from heel to toe (pointing forwards), while the z -axis is normal to the shoe's sole and pointing upwards (toward the ankle and leg of the wearer). Lastly, the y -axis is set in accordance to x and z , such that the frame is right-handed.

2.1.3. Data acquisition and notation

Data acquisition for \mathcal{G}_1 The participants of the first group \mathcal{G}_1 were instructed to wear the WGAS and walk with a self-selected gait on a treadmill. They were guided to complete three trials without assistance (e.g., handrail or body weight support) at $N_v = 3$ different walking speeds, specifically 0.53, 0.86, and 1.11 m/s. The treadmill (Bertec) was equipped with two force plates beneath the treadmill belt measuring ground reaction force sampled at 600 Hz. Retroreflective markers (marker size 14 mm) were placed on participants' shoes at the toes, the first and fifth Metatarsophalangeal joints, and the calcaneal tuberosities, as well as on the medial and lateral malleolus of the ankle, and on their hands. The markers on the hands were used to identify the hand clapping movements performed by the participants, which were used to signal the start of steady-state walking in the experiment. Positional data of markers were

collected by the motion capture system (Qualisys Opus 400 with 9 cameras) running at 100 Hz.

The data-acquisition protocol involved the following stages:

1. Data-logging was enabled for both systems, the WGAS and the motion-capture and force-plate system.
2. A simple calibration process was carried out before each walking trial began, consisting of turning-on and -off of the vibration on both shoes to give a signal for data synchronization. The vibration was registered by the WGAS as a binary signal (ON/OFF) and by the treadmill's force-plates. The registered vibration signals allowed for synchronization of the motion capture data and the WGAS data in post-processing.
3. The treadmill was subsequently ramped up to the respective speed and the participants were asked to acquaint themselves with the speed.
4. Once the treadmill reached the correct speed and the participants were comfortable walking at its speed, they performed a pronounced clapping motion with their hands to indicate the start of the steady-state motion capture measurement.
5. Participants then walked for approximately $N_s = 50$ strides for left and right foot, respectively. Stride count was monitored by a dedicated person.
6. After completing 50 strides on each side, the motion capture measurement was terminated, followed by a stop of the WGAS logging and a halting of the treadmill.

This protocol was repeated for each participant and speed.

Data obtained from the motion capture system contained the markers' positional data in three orthogonal axes, and the magnitude of the exerted force on the treadmill including its orientation vector in three axes. Data obtained from the WGAS consisted of acceleration, angular velocity and magnetic field vectors in three axes.

Data acquisition for \mathcal{G}_2 Each individual of the second group of participants $\mathcal{G}_2 = \{S_5, \dots, S_8\}$, was instructed to wear the WGAS and collect three distinct datasets. The first was a 2-min recording of their normal gait at their preferred walking speed. The second dataset consisted of the participant imitating various walking impairments, including spastic gait, steppage gait and waddling gait for another 2 min. Thirdly, they were asked to collect any range of non-stride motions, such as stomping, jumping, backwards and sideways walking, etc. The three data recordings were collected outdoors.

Notation Let $\mathbf{a}_x \in \mathbb{R}^m$ denote a column vector consisting of the sensor's acceleration data in x axis, where $\mathbf{a}_x[k]$ denotes the k -th entry of \mathbf{a}_x and is the acceleration data at time stamp k , $m \in \mathbb{N}$ denotes the length of the acceleration data, i.e., the number of collected samples. Similarly, \mathbf{a}_y and \mathbf{a}_z are column vectors of the sensor's acceleration data in y and z axis, respectively. Analogously, $\boldsymbol{\omega}_x$, $\boldsymbol{\omega}_y$, and $\boldsymbol{\omega}_z$ denote the data column vectors

¹ <https://www.freertos.org>

of angular velocity. Given any data vector, here we use the accelerometer signal \mathbf{a}_x as an example, $\mathbf{a}_x \in \mathbb{R}^m$, $\bar{\mathbb{E}}(\mathbf{a}_x) = \sum_{k=1}^m \frac{\mathbf{a}_x[k]}{m}$ denotes its sample mean. Time derivative signals of \mathbf{a}_x are denoted and computed as $\dot{\mathbf{a}}_x[k] = \frac{\mathbf{a}_x[k] - \mathbf{a}_x[k-1]}{t[k] - t[k-1]}$, where $\dot{\mathbf{a}}_x[0] = 0$ and \mathbf{t} is the vector of the timestamps. Similarly, the derivative signals, $\dot{\mathbf{a}}_y$, $\dot{\mathbf{a}}_z$, $\dot{\omega}_x$, $\dot{\omega}_y$, and $\dot{\omega}_z$, are computed in the same way. Given any real number $\alpha \in \mathbb{R}$, the absolute value of α is denoted as $|\alpha|$. Given any column vector \mathbf{a}_x , the symbol $|\mathbf{a}_x|$ denotes a new column vector obtained from the element-wise absolute value of \mathbf{a}_x . When applied to a set \mathcal{S} , the operator $|\cdot|$ indicates the cardinality of the set, i.e., the number of its elements $|\mathcal{S}| \in \mathbb{N}$.

2.2. Human labeling

2.2.1. Labeling strategy

The uncertainty of human gait labeling, which is widely accepted as the gold standard, is first evaluated in this paper. The datasets collected on the treadmill were labeled by $N = 5$ labelers ($\mathcal{G}_3 = \{S_9, \dots, S_{13}\}$). The uncertainty of manual labeling was quantified by studying the inter-labeler differences. The labelers were trained to use Mokka 3D Motion Kinematic & Kinetic Analyzer version 0.6.2² (Barre and Armand, 2014) to identify four ($N_e = 4$) key gait events, i.e., HO, TO, HS, and FF for both left and right feet. Note that different gait terminologies are used in different works (Vu et al., 2020). In this work, the following definitions for the gait events within a stride are applied: HS is defined as the first sample at which consistent contact between the heel and the ground is achieved after the swing phase; FF is defined as the first sample at which the tip of the foot is flat on the ground after an HS event; HO is defined as the first sample at which the heel (marker) leaves the ground after an FF event; and TO is defined as the first sample at which the toe (marker) leaves the ground after a HO event.

All labelers were given the same task of labeling 12 datasets collected from four ($N_h = 4$) healthy participants $S_1 - S_4$, each participant walking at three walking speeds (1.11, 0.86, and 0.53 m/s). This corresponds to roughly $N_L = 2 \cdot N_s \cdot N_v \cdot N_h \cdot N_e = 4,800$ events manually labeled by each labeler, leading to $N_E = N \cdot N_L = 24,000$ labeled events obtained in total. The labels obtained from each labeler are the motion capture system timestamps of the occurrence of each gait event.

2.2.2. Labeling condensation process

In general, labeling is not a difficult task for a human being, while it comes with some caveats, such as lossy marker data when the markers are disappeared for a few samples in the monitor due to marker occlusion, the lack of consistency of a

labeler³, and the lack of consistency among different labelers⁴. Labeling uncertainty is thus to be expected, which introduces a new problem: one cannot simply compute statistics on the manual events by constructing a matrix with the labels' arrays by all labelers since

1. there is no *a priori* guarantee that all labelers label the same number of gait events on a specific dataset;
2. even if all labelers provide the same number of labels, it does not mean that the obtained labels correspond to the same gait events. Figure 2A illustrates an example of this situation: Even if the two label sequences $(x_i)_{i=1}^7$ and $(y_j)_{j=1}^7$ from two labelers for the same dataset contain the same number of labels, the fourth (x_4, y_4) , fifth (x_5, y_5) , and sixth (x_6, y_6) pairs of labels do not correspond to the same gait events.

To evaluate the uncertainty in the manual labels for a specific gait event, each label of a labeler S_i needs to be first matched to the corresponding label of another labeler S_j . This is necessary for each pair of distinct labelers (S_i, S_j) for $(i, j) \in \{9, \dots, 13\} \times \{9, \dots, 13\}, i \neq j$. To achieve this labels-condensation, a matching algorithm is applied iteratively among pairs of labelers until the number of labels for each labeler converges to the same number and the labels from different labelers correspond to the same sequence of gait events. This matching algorithm is conducted for each gait event, side (left or right), and experiment (walking speed and walking participant) individually.

The employed matching algorithm is the Needleman-Wunsch algorithm (NWA) (Needleman and Wunsch, 1970), which was developed for DNA sequence matching. The basic idea of NWA is to apply dynamic programming to find the optimal alignment of two label sequences while allowing for the insertion/deletion of entries from either label sequence. The NWA is implemented in Python3 with `numpy`. To apply the NWA, given any two label sequences $(x_i)_{i=1}^m$ and $(y_j)_{j=1}^n$ for some $m, n \in \mathbb{N}$ from two labelers, the scoring function, including the match score (which measures the difference between two labels at the same index) and insertion/deletion score (which penalizes one insertion/deletion of a label in one of the label sequence), for the labeling condensation is defined as follows:

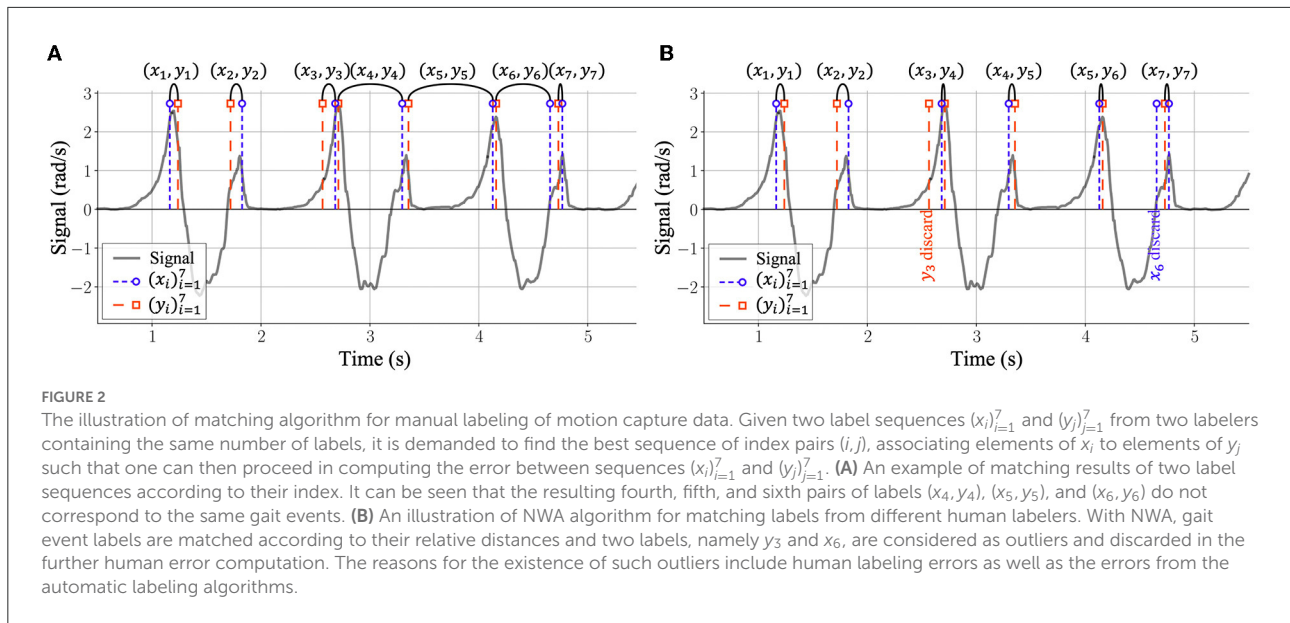
- Match score: $s(x_i, y_j) = |x_i - y_j|$;
- Insertion/deletion score: $s(x_i, y_j) = \bar{\mathbb{E}}((\delta x_i)_{i=1}^m)$, where $\delta x_i = x_i - x_{i-1}$ with $x_0 = 0$.

With this scoring function, two label sequences of gait events $(x_i)_{i=1}^m$ and $(y_j)_{j=1}^n$ are optimally aligned according to NWA algorithm (Needleman and Wunsch, 1970), leading to two

³ Even if the same labeler labels the same dataset several times in the same way, a different labeling sequence will be obtained each time.

⁴ Even if different labelers label the same dataset with the same labeling rules, different labeling sequences will be obtained for each labeler.

² <https://github.com/Biomechanical-Toolkit/Mokka>



sequences that have the same number of labels. An illustration of the NWA matching results is shown in Figure 2B.

For every gait event, walking speed, and walking participant, the corresponding five label sequences from five labelers were matched between each other using the NWA. After this matching process, the resulting five labeler event sequences have the same number of labels can be grouped in matrices for error analysis.

2.2.3. Limit of agreement

Given the matched gait event labels among five labelers, the uncertainty of the human labeling is quantified as the limit of agreement (LOA) and is assessed as follows. Let the vector I_i consist of all matched labels (labels from all gait speeds and walking participants datasets) of a gait event from a labeler $S_i \in \mathcal{G}_1$, and let $N = |\mathcal{G}_1|$. A type of gait event's label uncertainty is determined by first grouping all the labels of this type of gait event across the five labelers into a matrix as:

$$L = (I_1, I_2, \dots, I_N) \in \mathbb{R}^{M \times N} \quad (1)$$

where M is the number of matched labels for this type of gait event and N is the number of labelers. Next, the ground-truth labels, denoted as \bar{l} , of this type of gait event are computed by averaging the rows of L , i.e., $\bar{l} = \frac{1}{N} \sum_{i=1}^N I_i$. Then, an error matrix $D \in \mathbb{R}^{M \times N}$ is computed as $D[:, i] = \bar{l} - L[:, i]$, where $i \in \{1, 2, \dots, N\}$. With the obtained matrix D , the LOA is computed as the 95-th percentile of the ordered list (sorted from least to greatest) obtained from the absolute values of the entries in D . The ground-truth labels \bar{l} are considered as a benchmark for the further validation of the proposed gait event detection algorithm.

2.3. Automated gait analysis

This section describes the details of the proposed gait analysis algorithm. The flowchart of the algorithm is shown in Figure 3.

2.3.1. Data segmentation

The potential strides are first identified within a WGAS recording by analyzing the recorded raw signals of the angular rate and acceleration. To do so, the raw signals are first segmented into multiple regions of motion (ROMs), which represent continuous foot movements and potential strides, by applying filters and dynamic thresholds. The details of the applied filters and thresholds can be found in the Supplementary material.

2.3.2. Feature extraction

To classify whether an obtained ROM constitutes a stride or not, its underlying raw signals $\mathbf{R} = (\mathbf{a}_x, \mathbf{a}_y, \mathbf{a}_z, \boldsymbol{\omega}_x, \boldsymbol{\omega}_y, \boldsymbol{\omega}_z)$ are characterized by a set of features $\mathcal{F} = \{\mathbf{M}, \mathbf{c}\}$, consisting of a normalized and resampled IMU signal matrix, $\mathbf{M} \in \mathbb{R}^{100 \times 6}$, and a correlation score vector, $\mathbf{c} \in \mathbb{R}^{19}$. \mathbf{M} consists of six column vectors as $\mathbf{M} = (\hat{\mathbf{a}}_x, \hat{\mathbf{a}}_y, \hat{\mathbf{a}}_z, \hat{\boldsymbol{\omega}}_x, \hat{\boldsymbol{\omega}}_y, \hat{\boldsymbol{\omega}}_z)$, where given the acceleration vector \mathbf{a}_x within the considered ROM, the column vector $\hat{\mathbf{a}}_x$ is obtained by first min-max normalize \mathbf{a}_x to rescale its range into $[0, 1]$, and then resample the normalized data to 100 samples by interpolation using discrete Fourier transform (Bracewell and Bracewell, 1986). The obtained \mathbf{M} matrix is shown in Figure 3 as a grayscale image. The row vector \mathbf{c} quantifies the similarity between the considered ROM signal

\mathbf{R} and the template stride signal $\mathbf{T} = (\bar{a}_x, \bar{a}_y, \bar{a}_z, \bar{\omega}_x, \bar{\omega}_y, \bar{\omega}_z)$, where $\mathbf{T} \in \mathbb{R}^{100 \times 6}$. In this work, each column vector of \mathbf{T} is chosen as the average of 448 recorded stride signals, which have been resampled to 100 data points first, from four healthy participants $\mathcal{G}_2 = \{S_5, \dots, S_8\}$. The stride template signal \mathbf{T} reflects a normal stride in terms of acceleration and angular rate. To account for the temporal distortion of two signals \mathbf{R} and \mathbf{T} , which may be caused by different gait patterns such as walking speed, stride length, and cadence, a matching technique DTW (Berndt and Clifford, 1994; Giorgino et al., 2009) is applied to align each column of ROM signal \mathbf{R} onto each column of stride template signal \mathbf{T} . The rationale behind DTW is to stretch or compress the signal \mathbf{R} , such that the cumulative Euclidean distance, which is referred to as DTW distance, between the resulting warped signal $\tilde{\mathbf{R}}$ and \mathbf{T} is minimized under certain constraints. After \mathbf{R} has been warped to \mathbf{T} along the optimal warping path, the obtained DTW distance is normalized by dividing the optimal warping path length. The normalized DTW distance for i -th column of \mathbf{R} is denoted as d_i , where $i \in \{1, 2, \dots, 6\}$. In addition, the Pearson's correlation coefficient r_i and its p-value p_i for i -th column of the warped signal $\tilde{\mathbf{R}}$ and \mathbf{T} are calculated to characterize the similarity of \mathbf{R} and \mathbf{T} . Those p -values of Pearson's correlation are further rescaled by $\hat{p}_i = 0.1 \log p_i$. Together with the time duration of the considered ROM signal Δt , the correlation score vector \mathbf{c} is constructed as $\mathbf{c} = (\Delta t, \frac{1}{d_1}, r_1, \hat{p}_1, \dots, \frac{1}{d_6}, r_6, \hat{p}_6)^\top$.

2.3.3. Feature classification

The extracted features set \mathcal{F} of each ROM is then fed into a neural network, which classifies if the corresponding ROM signals represents a stride or not. The architecture of the proposed CNN, *stepperNet*, is illustrated in Figure 3, which accepts $\mathcal{F} = \{\mathbf{M}, \mathbf{c}\}$ as an input and outputs the probability of the feature set being a stride. It is composed of six total layers. The first is a convolutional layer, in which three kernels (with dimensions of 10×6), a step size of 1, and zero padding are applied to \mathbf{M} . The resulting layer ($91 \times 1 \times 3$) is applied with a max pooling kernel (3×1) to compress its dimensionality. The obtained three reduced feature maps are then concatenated into a single vector (90×1), to which \mathbf{c} is appended. The joined vector is then passed through two fully connected layers with 30 and 15 neurons each. Finally, the output of the network is normalized to a probability by the softmax function (Goodfellow et al., 2016). For any ROM signals: if the *stepperNet* outputs a probability greater than 0.5, the ROM is considered to be a walking stride, and thus the gait event detection algorithm will be further applied for intra-stride-based detection of gait events; if the stride probability is lower than 0.5, the ROM is assumed to be non-stride foot motion and will be rejected for further analysis.

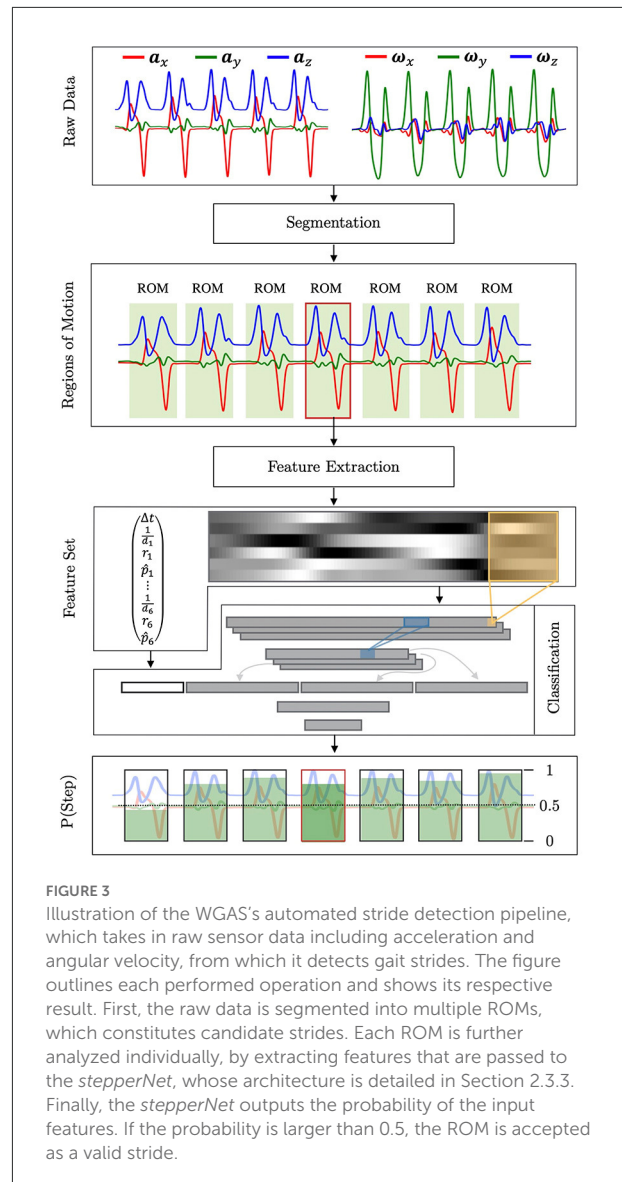
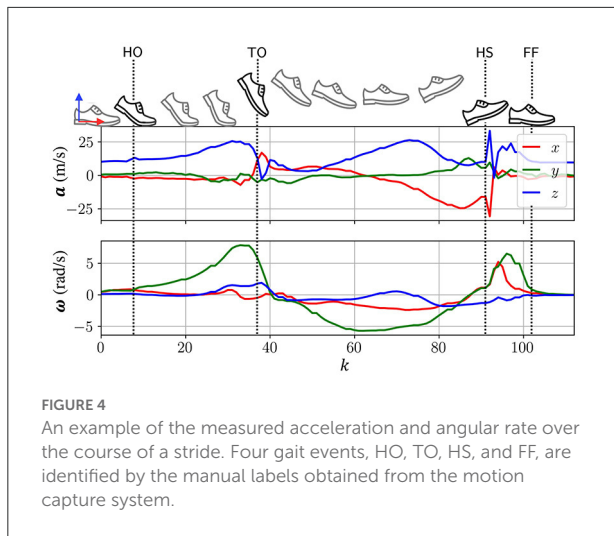


FIGURE 3 Illustration of the WGAS's automated stride detection pipeline, which takes in raw sensor data including acceleration and angular velocity, from which it detects gait strides. The figure outlines each performed operation and shows its respective result. First, the raw data is segmented into multiple ROMs, which constitutes candidate strides. Each ROM is further analyzed individually, by extracting features that are passed to the *stepperNet*, whose architecture is detailed in Section 2.3.3. Finally, the *stepperNet* outputs the probability of the input features. If the probability is larger than 0.5, the ROM is accepted as a valid stride.

2.3.4. Gait event detection

For any ROM, which has been determined to constitute a stride, the four gait events HO, TO, HS, and FF are identified by event-specific features in the recorded acceleration and angular rate within the signal region of the ROM. The correspondence of these features to the occurrence of the event was established experimentally by comparing the WGAS recordings to the manual labeled events of the accompanying motion capture system measurements obtained from Section 2.2. Figure 4 shows an example of the raw acceleration and angular rate over a stride recorded by the WGAS's IMU, together with the location of the gait events.

It can be observed that between the HO and FF, the foot can be assumed to be moving. As both gait events are the result of changes in the foot's orientation, in this



work, HO and FF are detected by taking the first and last time stamp where $\omega_a[k]/\sigma_{\omega_a} > 0.5$, where $\omega_a[k] = \sqrt{(\omega_x[k])^2 + (\omega_y[k])^2 + (\omega_z[k])^2}$ and σ_{ω_a} is the standard deviation of the column vector ω_a . As the HS event resembles an impulse that manifests itself in an abrupt change in the measured acceleration and its impact is markedly experienced in the z -direction, the HS event is determined by the location of the most prominent peak in \dot{a}_z . By comparing the WGAS's measurements to the manual labels of the motion capture system the TO event is determined as the first timestamp, where $a_x[k] = a_z[k]$.

2.3.5. Data synchronization

The validation dataset consists of two WGAS measurements (left and right) with a corresponding motion capture system recording, each of which is logged with its own internal clocks. To compare the manually labeled gait events to the events detected by the gait event detection algorithm, we first ensure that they are temporally synchronized. To this end, each WGAS is set to initiate a vibration sequence once they are turned on, vibrating for 1 s, a 1 s pause, and a second vibration of 1 s. As the experiment starts with the participant standing on the treadmill of the motion capture system the vibration sequence is also picked up by the treadmill's integrated force plate. This allows for the temporal synchronization between the data collected by the two systems.

The proposed algorithms of gait stride detection and gait event detection are implemented in Python3. Further important tools include PyTorch, which was used as the machine-learning framework (Paszke et al., 2019) and the dtw-python library as the DTW library (Giorgino et al., 2009). Lastly, the proposed algorithms extensively use various toolboxes of the SciPy library (Virtanen et al., 2020).

TABLE 1 Labels matching results obtained by applying the NWA across the five labelers's manual labeling of motion capture system data.

Event	Labels	Removed	Percentage
HO	6,916	121	1.75
TO	6,930	95	1.37
HS	6,924	84	1.21
FF	6,913	113	1.63
Total	27,683	413	1.49

3. Results

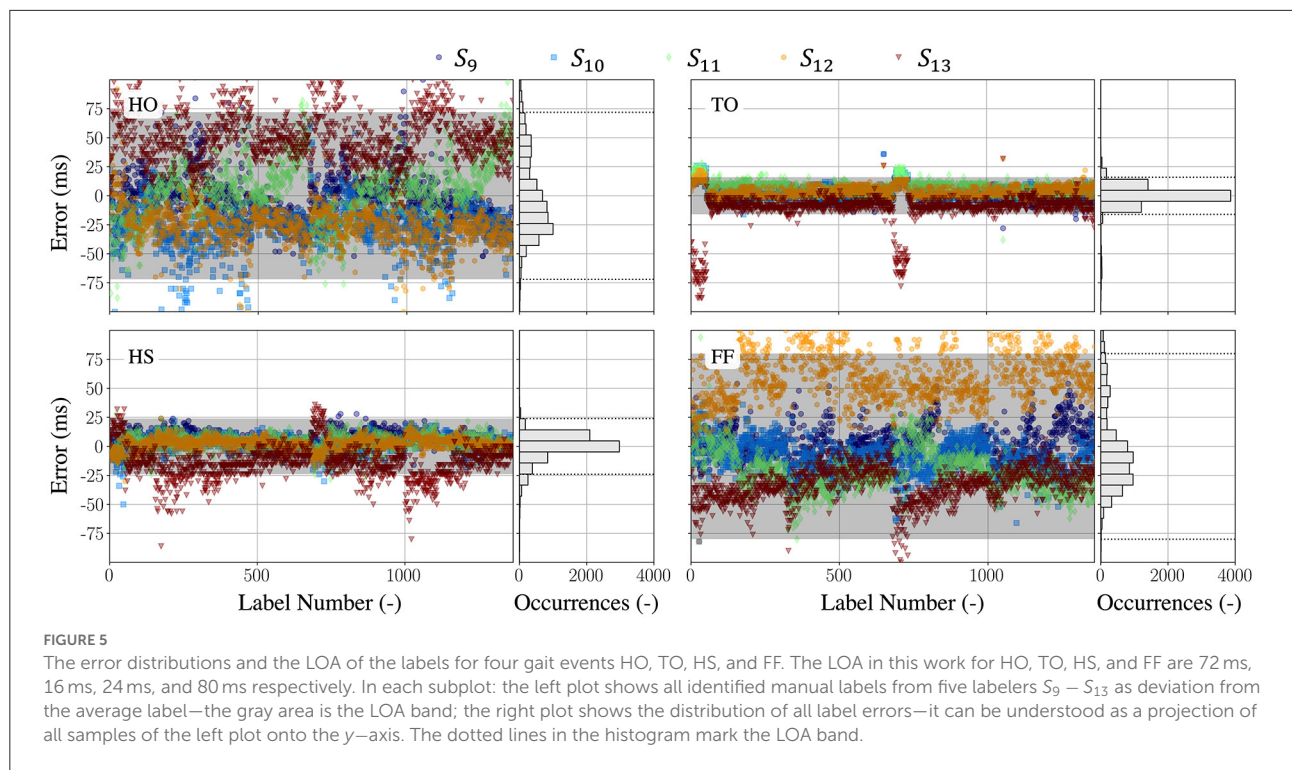
3.1. Uncertainty of human labeling

For all labeled datasets, the label matching results of the NWA algorithm are presented in Table 1. It can be seen that only a small percentage (1.49%) of labels were discarded for all gait events. Overall, 413 labels were discarded from a total of 27,683 manually labeled events. The maximum number of discarded events was 23 out of 333 available (for HO event of participant S_{12} with walking speed 1.11 m/s).

After the procedure of label condensation, the LOA results illustrated in Figure 5 are obtained. The LOA in this work for HO, TO, HS, and FF are 72, 16, 24, and 80 ms, respectively, which corresponds to 2–8 samples of uncertainty/disagreement when sampling at 100 Hz. It can be seen that for the main gait events HS and TO, the labels agree well among labelers as the LOA are 16 and 20 ms, which means that only two samples on average differ between the five labelers. However, the human labeling uncertainty for the secondary gait events HO and FF reached 72 and 80 ms, i.e., nearly eight samples of variations are observed between the labels labeled by the five labelers, which are higher than the LOA of the HS and TO labels. This is due to the intrinsic difficulty that HO and FF come within their definitions, and consequent difficulties for humans to decide from the motion capture system data whether an event of these types has occurred. As a consequence, the limits in terms of accuracy claims of automated systems are to be set around the identified LOA, i.e., the LOA is the limit of accuracy for the validated system as it represents the intrinsic uncertainty in the ground-truth measurements.

3.2. Gait cycle segmentation

The second dataset collected for building *stepperNet* consisted of 4,014 labeled feature sets, of which 1,888 represented strides, and the rest characterized non-stride foot movement. For training *stepperNet*, the collected data were randomly split into a training dataset (3,493 features), with which the network's parameters were optimized, and a test

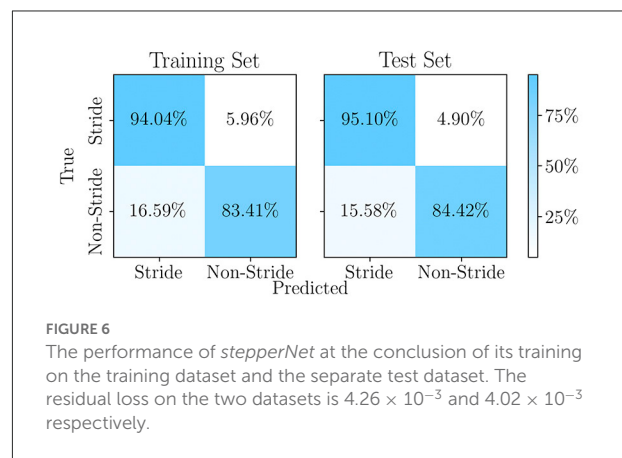


set (521 features) which was used to monitor *stepperNet*'s performance in training.

The *stepperNet* was trained by stochastic gradient descent (Bottou, 2010) over 7 epochs, with the learning rate and momentum set to 0.015 and 0.75, respectively. Due to the small network size and the number of training examples, the performance of the trained network is highly dependent on the initially chosen weights and biases. To optimize the stride classification capabilities of *stepperNet*, the training procedure was repeated multiple times, each time with different random initial internal parameters, and the highest scoring network was saved and deployed. The optimized *stepperNet*'s performance on the training and validation data are reported in Figure 6, with stride and non-stride detection accuracy of 95.10% and 84.42%, respectively.

3.3. Gait event detection

To evaluate the performance of the gait event detection algorithm, the timestamp of each type of gait event detected by the proposed algorithm is compared to the ground truth labels, i.e., the timestamp of the gait event identified manually in the motion capture system data. Note that the ground truth labels obtained from Section 2.2 are served as the baseline for the validation of this gait event detection algorithm. The performance of the gait event detection algorithm is quantified by four error metrics for each type of gait event, i.e., the mean



error, the mean absolute error (MAE), the root mean squared error (RMSE), and the 95-th percentile error (LOA). The overall error metrics for the detected HO, TO, HS and FF, averaged across all participants and treadmill speeds, are presented in Table 2. In total, the timestamps of the WGAS detected gait events for about 1,350 strides are compared to the ground truth labels. This number has a slight differences for different gait events as shown in Table 2, this is due to the reason that the manual labels of a gait event that are not identified by all labelers has been discarded in the labeling condensation process.

The error metrics are also presented for each event averaged across all participants but grouped by treadmill speed as shown

TABLE 2 Error metrics for WGAS gait event detection results.

Event	Mean (ms)	STD (ms)	MAE (ms)	RMSE (ms)	LOA (ms)	Samples (-)
HO	9	29	24	31	60	1,338
TO	2	7	5	7	10	1,347
HS	-2	16	9	16	20	1,355
FF	9	34	13	36	30	1,349

in Figure 7 and Supplementary Table S1. It can be observed that as the walking speed increases from 0.53 to 1.11 m/s, the event detection error of the algorithm appears to slightly decrease (HO: from 11 ± 22 ms to 8 ± 27 ms, TO: from 3 ± 7 ms to 1 ± 6 ms, HS: from -5 ± 14 ms to -4 ± 13 ms, FF: from 15 ± 17 ms to 6 ± 12 ms). This can be explained by the reason that with the higher walking speed, the event features are more pronounced and unambiguous within the signals. As this effect is negligible, the gait event detection algorithm can be considered invariant and robust toward changes in gait speed.

Generally, it should be stated that it is difficult to discern if an error is the result of faulty event detection of the algorithm or due to the subjective nature of manual labeling. For the HO and FF, this is especially true, as the events are not definitively identifiable in the motion capture system data by eye introducing an error in the ground truth itself. An infinite number of labelers may average out every individual labeling error, but due to time and cost constraints, the limited number of labelers in this study and in practice, the ground truth may still be imperfect.

4. Discussion

Manual labeling of four gait events HO, TO, HS, and FF on motion capture system datasets are evaluated in this work as they are essential for analyzing temporal gait parameters, such as stride time, swing time, stance time, and single- and double-support time. As can be seen from Figure 5, The manual labels that were labeled by five labelers with the similar labeling experience exhibit a remarkable consistency among different labelers for two main gait events, HS and TO, with only two samples (around 20 ms) of variations are observed. However, they showed greater disagreement on the secondary gait events HO and FF, with up to eight samples of average variations are observed on these two gait events. This disagreement is due to the inherent challenges of observing HO and FF in optical data from motion capture systems and the fact that different labelers have different perspectives even though they have similar labeling experience. It can also be observed from the Figure 5 that even for the same labeler, such as S13, they have a large labeling deviation along the sequence of data labeling, varying from 2 to 8 samples. It is the intrinsic human labeling uncertainty regardless of human experience of labeling. This

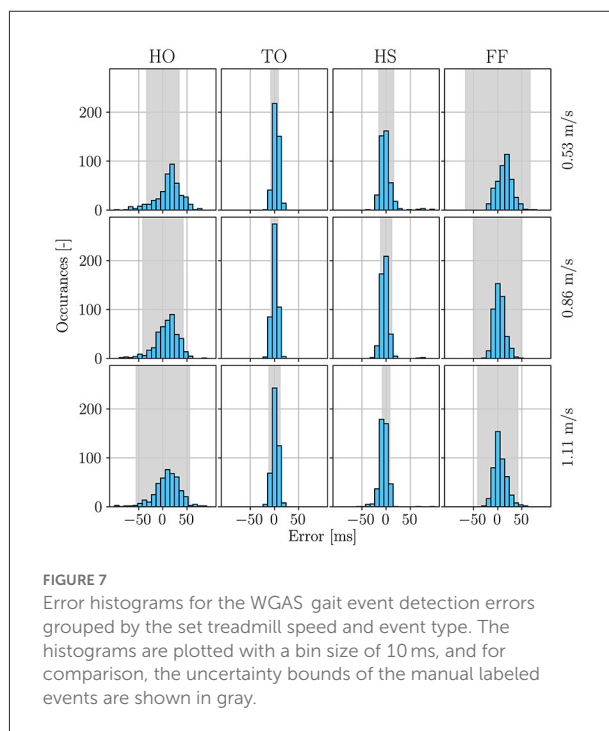


FIGURE 7 Error histograms for the WGAS gait event detection errors grouped by the set treadmill speed and event type. The histograms are plotted with a bin size of 10 ms, and for comparison, the uncertainty bounds of the manual labeled events are shown in gray.

uncertainty represents the accuracy limit of human gait-labeling based on motion capture data, and it determines the limit for any validation study which considers the motion capture system as the gold standard. Low consistency between labelers can have a direct impact on gait event detection reliability. This further affects the estimation accuracy of spatio-temporal gait parameters, as well as the medical decision-making and therapy by clinicians.

The reasons for identifying HS, TO, HO, and FF gait events in a gait cycle are discussed as follows. The gait cycles can be delimited by the HS event, which is considered of paramount functional importance marking the shift from flexors to anti-gravity muscles and is also reliably detectable (Burnfield, 2010). To calculate gait parameters for each gait cycle, the movement data are interpolated from HS to the next ipsilateral HS. Errors in HS detection lead to normalization artifacts that can affect the whole gait cycle and hence all calculated parameters. TO delineates the two main phases of gait: stance and swing. Kinematic parameters are frequently reported for these phases separately, hence errors in TO detection affect not only values

at TO but also parameters calculated for the phases, e.g., “knee range of motion during swing.” At both time points, parameters are frequently reported that are in phases of rapid change, such as sagittal ankle and knee angle at TO (both in the range of 100–300°/s in dependence on walking speed), or foot elevation angle at HS (Mentiplay et al., 2018). Concerning the gait events HO and FF, there is a growing interest in the importance of these phases that depict a first reaction to a state change from swing-to-stance (Nolan and Yarossi, 2011) and the preparatory phase from stance-to-swing (Lewek et al., 2018; Roelker et al., 2019). Detecting these additional gait events also allows the calculation of parameters related to push off and initial contact. These events do not correspond with traditional clinical phases of gait, however present a technical method to acquire information about critical states. Overall, more detailed segmentation of the gait cycle and calculation of movement parameters that provide a relationship to typical recovery patterns in clinical populations allows clinicians to make more objective and detailed assessments of the deviation in pathological gait and the improvements in rehabilitation progress. Furthermore, gait event detection is also critical for applications in which gait events are served to trigger assistive devices, as well as when considering the use of orthotic or therapeutic interventions, particularly in functional electrical stimulation (Zahradka et al., 2020).

After the manual labels obtained on the motion capture data, a *stepperNet* to automatically identify strides and reject non-strides from long-term walking signals is developed, and a gait event detection algorithm was adopted to segment the identified strides into sub-phases for further estimation of gait parameters. As shown in Figure 6, the stride classifier *stepperNet* shows a good classification performance with an accuracy of 94.04% in stride detection. For non-stride detection, a relatively low accuracy of 83.41% was obtained. This could be due to the fact that in the motion activity of healthy people, stride motion include only walking movements, while non-stride motion include various human motions, such as jumping, sitting and standing, climbing, etc. The variability of stride signals are much higher than the variability of non-stride signals. It can also be noticed that the classification accuracy of the test set is slightly higher than that of the training set (around 1%), this can be explained by the randomness when splitting the total feature set, which yielded marginally 1% more difficult examples in the training set than in the test set. Note that the CNN is only used for identifying a valid stride and eliminating non-stride data from the dataset that was acquired over a long (walking) period. It shall be noted that whilst one is walking for longer periods in real-life scenarios not all motions are actually strides, e.g., walking usually features breaks and motions associated to breaks such as pivoting on the spot, which are not to be considered strides/walking. The presented hybrid model allows to split the long-term “walking” data into “stride” and “non-stride” data slices, which makes further data processing (gait event) easier.

After stride segmentation, gait events are afterwards detected within the region of this identified stride by searching for specific signal characteristics. The detected gait events are validated with the ground-truth labels obtained from the manual labeling of motion capture system data. It can be seen from Table 2, the gait event detection algorithm perform well for all four gait events (HO, TO, HS, and FF), with a mean difference of less than one samples and a standard deviation of less than four samples compared to the benchmark (i.e., the labels from the motion capture system). The errors of gait events HO and FF are slightly higher than those of TO and HS. This is due to the reason that at the gait event of HO and FF (Figure 4), the signals are flat, the amplitude or the rate fluctuations of the motion signal are not as pronounced as in the TO and HS. This agrees with the results of human labeling in Figure 5, where the uncertainties of HO and FF are higher than that of TO and HS. Besides, comparing the LOA of the detected gait events, as shown in Table 2, with those of human labeling in Figure 5, it can be seen that the gait event detection algorithm are generally consistent in their predictions. Table 3 compares the results of this paper to the work reported in the literature in terms of MAE. This comparison focuses on studies with healthy participants, no pathological patients were involved in the experiments. The comparison shows that the proposed algorithm has a comparable performance in detecting HS and TO. TO has the lowest MAE compared to the other studies, while the MAE of HS is slightly higher MAE than the study of Fadillioglu et al. (2020) and Mo and Chow (2018). However, the MAE of HS is less than one sampling time (10 ms), this difference is still within the acceptable range.

Opposed to the commonly used approach in the literature, which straightly scans the entire walking recording to identify the gait events and then employs the gait event to define a stride, the CNN eliminates all non-stride data and only keeps stride data for further gait event detection. The main advantage of this approach is robustness, as the aforementioned method scans the entire dataset, erroneously detected gait events, have the potential to invalidate the entire analysis. Errors in our proposed method however only affect an individual stride and the detection of gait events belonging to other strides is undisturbed.

This work is based on a limited number and type of validation participants, but it still provides important insight on analyzing gait in a robust and efficient way. Furthermore, the selection of manual labeling of the motion capture data as the ground truth induces caveats. Vertical ground reaction force, specifically, was not used in this study due to a large amount of foot strikes that loaded both force plates. These events would have needed to be excluded, which would have introduced a bias into the data set. Therefore, group-based manual labeling was performed, which is also representative of how pathological gait is frequently segmented. Though males and females differ from a biomechanical perspective due to their

TABLE 3 A comparison of the MAE for HS and TO gait event detection in healthy participants with those reported in the literature.

Study	Sensor location	Walking condition	HS (ms)	TO (ms)
Tiwari and Joshi (2020)	Foot	Over-ground walking	22.87	10.42
Fadillioğlu et al. (2020)	Shank	Treadmill walking	7	9
Flood et al. (2019)	Waist and shank	Treadmill walking	16.41	39.84
Nazmi et al. (2019)	Foot	Treadmill walking	35	49
Mo and Chow (2018)	Pelvis	10-m walkway walking	6.2	20.3
Godiyal et al. (2018)	Thigh	Over-ground walking	9.66	16.99
Ours	Foot	Treadmill walking	9	5

differences in joints-relative position, the proposed system is believed to be agnostic with regard to these differences. This is because that different joint positions and body dimensions affect the motion metrics at the extremities, which is where the proposed system performs the measurements. Therefore, the motion captured by the WGAS system is not subject to potential gender-related effects. While the issue of gender balance and pathological gait have to be addressed in the future for assessing the generalizability and effectiveness of the proposed method for pathological subjects. In this study, the stride template was generated by averaging a set of healthy strides. For pathological gait, strides will not be as homogeneous, new methodologies have to be developed for optimizing the stride template, such as generating template from kernel density estimation or Gaussian mixture model, constructing a library of stride templates from different pathologies or experimental conditions. In this work, *stepperNet* is only validated on treadmill walking data, one participant's stride is likely to be identical during the same speed session. As the *stepperNet* is primarily designed for identifying walking strides and eliminating the non-walking strides data from long-term free walking data to reduce the computational cost, the performance of *stepperNet* on long-term free walking in the real-world must be validated. Future work will focus on investigating the detection accuracy of *stepperNet* in various conditions, including variable speed walking, free-range walking, and various clinical populations.

5. Conclusion

We described a method to estimate and analyze the uncertainty in manual labeling of human walking data from motion capture systems, which are commonly used as a benchmark. Human manual labeling uncertainty was assessed by measuring inter-labeler inconsistencies. The limits of detection for key gait events, HO, TO, HS, and FF have been estimated as 72, 16, 24, and 80 ms respectively with a sampling time of 10 ms. Those inherent label uncertainty of human gait labeling based on motion capture data is the accuracy limit for any validation research using this technology.

Using the manual labels (obtained by averaging across all labelers) as a baseline, we also present a novel CNN-based algorithm capable of detecting key gait events with a detection error less than 25 ms in healthy, fixed speed treadmill walking. The algorithm provides an improvement on current solutions in that it allows the accurate processing of long continuous walking data streams. Clinically, the accurate detection of these events reduce distortion effects when calculating gait parameters, providing a gateway to reliable parameter calculation that are lower than the minimal clinically important difference. Achieving a high accuracy in gait parameters is paramount for empowering clinical decision-making based on mobile gait analysis.

Data availability statement

The original contributions presented in the study are included in the article/[Supplementary material](#), further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Swiss Ethics Committee (BASEC Nr Req-2019-00715) following the Good Clinical Practice Guidelines and the Declaration of Helsinki from. The patients/participants provided their written informed consent to participate in this study.

Author contributions

JW and AS conceived and designed the study. JW, HM, AS, and BB performed the experiment and analysis. CA and GC collected the data. JW took the lead in writing the manuscript with input from HM. AS, BB, and JW was in charge of overall planning. OE, SP, and BN acquired the funding of the study. All authors provided critical feedback and helped shape the manuscript.

Funding

This project has received funding from the European Union's Horizon 2020 Research and Innovation program under the Marie Skłodowska-Curie grant agreement no. 764977. Open access funding provided by ETH Zurich.

Conflict of interest

Authors JW, HM, AS, BB, GC, and OE were employed by Magnes AG. Authors BN and SP are shareholders in Magnes AG.

The remaining author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

References

- Agostini, V., Balestra, G., and Knaflitz, M. (2013). Segmentation and classification of gait cycles. *IEEE Trans. Neural Syst. Rehabil. Eng.* 22, 946–952. doi: 10.1109/TNSRE.2013.2291907
- Atrsaei, A., Corrà, M. F., Dadashi, F., Vila-Chã, N., Maia, L., Mariani, B., et al. (2021). Gait speed in clinical and daily living assessments in parkinson's disease patients: performance versus capacity. *NPJ Parkinsons Dis.* 7, 1–11. doi: 10.1038/s41531-021-00171-0
- Barre, A., and Armand, S. (2014). Biomechanical toolkit: open-source framework to visualize and process biomechanical data. *Comput. Methods Programs Biomed.* 114, 80–87. doi: 10.1016/j.cmpb.2014.01.012
- Barth, J., Oberndorfer, C., Pasluosta, C., Schüle, S., Gassner, H., Reinfelder, S., et al. (2015). Stride segmentation during free walk movements using multi-dimensional subsequent dynamic time warping on inertial sensor data. *Sensors* 15, 6419–6440. doi: 10.3390/s150306419
- Berndt, D. J., and Clifford, J. (1994). "Using dynamic time warping to find patterns in time series," in *KDD Workshop, Vol. 10* (Seattle, WA), 359–370.
- Bottou, L. (2010). "Large-scale machine learning with stochastic gradient descent," in *Proceedings of COMPSTAT'2010* (Springer: Paris), 177–186.
- Bracewell, R. N., and Bracewell, R. N. (1986). *The Fourier Transform and Its Applications, Vol. 31999*. New York, NY: McGraw-Hill.
- Brogna, L., Palumbo, P., Grimm, B., and Palmerini, L. (2019). Assessing gait in Parkinson's disease using wearable motion sensors: a systematic review. *Diseases* 7, 18. doi: 10.3390/diseases7010018
- Burnfield, M. (2010). Gait analysis: normal and pathological function. *J. Sports Sci. Med.* 9, 353.
- Caldas, R., Mundt, M., Potthast, W., de Lima Neto, F. B., and Markert, B. (2017). A systematic review of gait analysis methods based on inertial sensors and adaptive algorithms. *Gait Posture* 57, 204–210. doi: 10.1016/j.gaitpost.2017.06.019
- Carse, B., Meadows, B., Bowers, R., and Rowe, P. (2013). Affordable clinical gait analysis: an assessment of the marker tracking accuracy of a new low-cost optical 3d motion analysis system. *Physiotherapy* 99, 347–351. doi: 10.1016/j.physio.2013.03.001
- Celik, Y., Stuart, S., Woo, W. L., Sejdic, E., and Godfrey, A. (2022). Multi-modal gait: a wearable, algorithm and data fusion approach for clinical and free-living assessment. *Inf. Fusion* 78, 57–70. doi: 10.1016/j.inffus.2021.09.016
- Del Din, S., Godfrey, A., Galna, B., Lord, S., and Rochester, L. (2016). Free-living gait characteristics in ageing and parkinson's disease: impact of environment and ambulatory bout length. *J. Neuroeng. Rehabil.* 13, 1–12. doi: 10.1186/s12984-016-0154-5
- DeLisa, J. A. (1998). *Gait Analysis in the Science of Rehabilitation, Vol. 2*. Diane Publishing.
- Eng, J. J., and Tang, P.-F. (2007). Gait training strategies to optimize walking ability in people with stroke: a synthesis of the evidence. *Expert. Rev. Neurother.* 7, 1417–1436. doi: 10.1586/14737175.7.10.1417
- Fadillioglu, C., Stetter, B. J., Ringhof, S., Krafft, F. C., Sell, S., and Stein, T. (2020). Automated gait event detection for a variety of locomotion tasks using a novel gyroscope-based algorithm. *Gait Posture* 81, 102–108. doi: 10.1016/j.gaitpost.2020.06.019
- Flood, M. W., O'Callaghan, B. P., and Lowery, M. M. (2019). Gait event detection from accelerometry using the teager-kaiser energy operator. *IEEE Trans. Biomed. Eng.* 67, 658–666. doi: 10.1109/TBME.2019.2919394
- Giorgino, T., et al. (2009). Computing and visualizing dynamic time warping alignments in R: the dtw package. *J. Stat. Softw.* 31, 1–24. doi: 10.18637/jss.v031.i07
- Godiyal, A. K., Verma, H. K., Khanna, N., and Joshi, D. (2018). A force myography-based system for gait event detection in overground and ramp walking. *IEEE Trans. Instrum. Meas.* 67, 2314–2323. doi: 10.1109/TIM.2018.2816799
- Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep Learning*. MIT Press.
- Gras, L. Z., Kanaan, S. F., McDowd, J. M., Colgrove, Y. M., Burns, J., and Pohl, P. S. (2015). Balance and gait of adults with very mild Alzheimer's disease. *J. Geriatr. Phys. Therapy* 38, 1. doi: 10.1519/JPT.0000000000000020
- Heijmans, M., Habets, J. G., Herff, C., Aarts, J., Stevens, A., Kuijff, M. L., et al. (2019). Monitoring parkinson's disease symptoms during daily life: a feasibility study. *NPJ Parkinsons Dis.* 5, 1–6. doi: 10.1038/s41531-019-0093-5
- Hickey, A., Del Din, S., Rochester, L., and Godfrey, A. (2016). Detecting free-living steps and walking bouts: validating an algorithm for macro gait analysis. *Physiol. Meas.* 38, N1. doi: 10.1088/1361-6579/38/1/N1
- Hoskovicová, M., Ulmanová, O., Šprdlík, O., Sieger, T., Nováková, J., Jech, R., et al. (2013). Disorders of balance and gait in essential tremor are associated with midline tremor and age. *Cerebellum* 12, 27–34. doi: 10.1007/s12311-012-0384-4
- Hsu, W.-C., Sugiarto, T., Lin, Y.-J., Yang, F.-C., Lin, Z.-Y., Sun, C.-T., et al. (2018). Multiple-wearable-sensor-based gait classification and analysis in patients with neurological disorders. *Sensors* 18, 3397. doi: 10.3390/s18103397
- Jacobi, H., Hauser, T.-K., Giunti, P., Globas, C., Bauer, P., Schmitz-Hübsch, T., et al. (2012). Spinocerebellar ataxia types 1, 2, 3 and 6: the clinical spectrum of ataxia and morphometric brainstem and cerebellar findings. *Cerebellum* 11, 155–166. doi: 10.1007/s12311-011-0292-z
- Kaye, J., Mattek, N., Dodge, H., Buracchio, T., Austin, D., Hagler, S., et al. (2012). One walk a year to 1000 within a year: continuous in-home unobtrusive gait assessment of older adults. *Gait Posture* 35, 197–202. doi: 10.1016/j.gaitpost.2011.09.006
- Kidziński, Ł., Delp, S., and Schwartz, M. (2019). Automatic real-time gait event detection in children using deep neural networks. *PLoS ONE* 14, e0211466. doi: 10.1371/journal.pone.0211466

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fnins.2022.976594/full#supplementary-material>

- Kobsar, D., Charlton, J. M., Tse, C. T., Esculier, J.-F., Graffos, A., Krowchuk, N. M., et al. (2020). Validity and reliability of wearable inertial sensors in healthy adult walking: a systematic review and meta-analysis. *J. Neuroeng. Rehabil.* 17, 1–21. doi: 10.1186/s12984-020-00685-3
- Laughman, R. K., Askew, L. J., Bleimeyer, R. R., and Chao, E. Y. (1984). Objective clinical evaluation of function: gait analysis. *Phys. Ther.* 64, 1839–1845. doi: 10.1093/ptj/64.12.1839
- Lees, A. (1992). When did ray Kennedy's Parkinson's disease begin? *Mov. Disord.* 7, 110–116. doi: 10.1002/mds.870070203
- Lempereur, M., Rousseau, F., Rémy-Néris, O., Pons, C., Houx, L., Quellec, G., et al. (2020). A new deep learning-based method for the detection of gait events in children with gait disorders: Proof-of-concept and concurrent validity. *J. Biomech.* 98, 109490. doi: 10.1016/j.jbiomech.2019.109490
- Lewek, M. D., Raiti, C., and Doty, A. (2018). The presence of a paretic propulsion reserve during gait in individuals following stroke. *Neurorehabil. Neural Repair.* 32, 1011–1019. doi: 10.1177/1545968318809920
- Liu, L., Wang, H., Li, H., Liu, J., Qiu, S., Zhao, H., et al. (2021). Ambulatory human gait phase detection using wearable inertial sensors and hidden markov model. *Sensors* 21, 1347. doi: 10.3390/s21041347
- Mancini, M., Bloem, B. R., Horak, F. B., Lewis, S. J., Nieuwboer, A., and Nonnekes, J. (2019). Clinical and methodological challenges for assessing freezing of gait: future perspectives. *Mov. Disord.* 34, 783–790. doi: 10.1002/mds.27709
- Martindale, C. F., Christlein, V., Klumpp, P., and Eskofier, B. M. (2021). Wearables-based multi-task gait and activity segmentation using recurrent neural networks. *Neurocomputing* 432, 250–261. doi: 10.1016/j.neucom.2020.08.079
- Mc Ardle, R., Galna, B., Donaghy, P., Thomas, A., and Rochester, L. (2019). Do Alzheimer's and lewy body disease have discrete pathological signatures of gait? *Alzheimers Dement.* 15, 1367–1377. doi: 10.1016/j.jalz.2019.06.4953
- Meng, X., Yu, H., and Tham, M. P. (2013). "Gait phase detection in able-bodied subjects and dementia patients," in *2013 35th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)* (Osaka: IEEE), 4907–4910.
- Mentiplay, B. F., Banky, M., Clark, R. A., Kahn, M. B., and Williams, G. (2018). Lower limb angular velocity during walking at various speeds. *Gait Posture* 65, 190–196. doi: 10.1016/j.gaitpost.2018.06.162
- Mihradi, S., Dirgantara, T., and Mahyuddin, A. I. (2011). "Development of an optical motion-capture system for 3D gait analysis," in *2011 2nd International Conference on Instrumentation, Communications, Information Technology, and Biomedical Engineering* (Bandung: IEEE), 391–394.
- Mo, S., and Chow, D. H. (2018). Accuracy of three methods in gait event detection during overground running. *Gait Posture* 59, 93–98. doi: 10.1016/j.gaitpost.2017.10.009
- Nazmi, N., Rahman, M. A. A., Yamamoto, S.-I., and Ahmad, S. A. (2019). Walking gait event detection based on electromyography signals using artificial neural network. *Biomed. Signal Process. Control* 47, 334–343. doi: 10.1016/j.bspc.2018.08.030
- Needleman, S. B., and Wunsch, C. D. (1970). A general method applicable to the search for similarities in the amino acid sequence of two proteins. *J. Mol. Biol.* 48, 443–453. doi: 10.1016/0022-2836(70)90057-4
- Nolan, K. J., and Yarossi, M. (2011). Preservation of the first rocker is related to increases in gait speed in individuals with hemiplegia and afo. *Clin. Biomech.* 26, 655–660. doi: 10.1016/j.clinbiomech.2011.03.011
- Nürnberg, L., Klein, C., Baudrexel, S., Roggendorf, J., Hildner, M., Chen, S., et al. (2015). Ultrasound-based motion analysis demonstrates bilateral arm hypokinesia during gait in heterozygous pink1 mutation carriers. *Mov. Disord.* 30, 386–392. doi: 10.1002/mds.26127
- Oudre, L., Barrois-Müller, R., Moreau, T., Truong, C., Vienne-Jumeau, A., Ricard, D., et al. (2018). Template-based step detection with inertial measurement units. *Sensors* 18, 4033. doi: 10.3390/s18114033
- Palmerini, L., Rocchi, L., Mazilu, S., Gazit, E., Hausdorff, J. M., and Chiari, L. (2017). Identification of characteristic motor patterns preceding freezing of gait in Parkinson's disease using wearable sensors. *Front. Neurol.* 8, 394. doi: 10.3389/fneur.2017.00394
- Papavasiliou, I., Zhang, W., and Han, S. (2017). Real-time data-driven gait phase detection using ground contact force measurements: algorithms, platform design and performance. *Smart Health* 1, 34–49. doi: 10.1016/j.smhl.2017.03.001
- Paszke, A., Gross, S., Massa, F., Lerer, A., Bradbury, J., Chanan, G., et al. (2019). Pytorch: an imperative style, high-performance deep learning library. *arXiv preprint arXiv:1912.01703*.
- Pistacchi, M., Gioulis, M., Sanson, F., De Giovannini, E., Filippi, G., Rossetto, F., et al. (2017). Gait analysis and clinical correlations in early Parkinson's disease. *Funct. Neurol.* 32, 28. doi: 10.11138/FNur/2017.32.1.028
- Pulliam, C. L., Heldman, D. A., Brokaw, E. B., Mera, T. O., Mari, Z. K., and Burack, M. A. (2017). Continuous assessment of levodopa response in Parkinson's disease using wearable motion sensors. *IEEE Trans. Biomed. Eng.* 65, 159–164. doi: 10.1109/TBME.2017.2697764
- Rehman, R. Z. U., Del Din, S., Guan, Y., Yarnall, A. J., Shi, J. Q., and Rochester, L. (2019). Selecting clinically relevant gait characteristics for classification of early parkinson's disease: a comprehensive machine learning approach. *Sci. Rep.* 9, 1–12. doi: 10.1038/s41598-019-53656-7
- Richards, C. L., Malouin, F., and Dean, C. (1999). Gait in stroke: assessment and rehabilitation. *Clin. Geriatr. Med.* 15, 833–856. doi: 10.1016/S0749-0690(18)0034-X
- Robles-García, V., Corral-Bergantiños, Y., Espinosa, N., Jácome, M. A., García-Sancho, C., Cudeiro, J., et al. (2015). Spatiotemporal gait patterns during overt and covert evaluation in patients with parkinson's disease and healthy subjects: is there a Hawthorne effect? *J. Appl. Biomech.* 31, 189–194. doi: 10.1123/jab.2013-0319
- Roelker, S. A., Bowden, M. G., Kautz, S. A., and Neptune, R. R. (2019). Paretic propulsion as a measure of walking performance and functional motor recovery post-stroke: a review. *Gait Posture* 68, 6–14. doi: 10.1016/j.gaitpost.2018.10.027
- Selves, C., Stoquart, G., and Lejeune, T. (2020). Gait rehabilitation after stroke: review of evidence of predictors, clinical outcomes and timing for interventions. *Acta Neurol Belg.* 120, 783–790. doi: 10.1007/s13760-020-01320-7
- Senanayake, C. M., and Senanayake, S. A. (2010). Computational intelligent gait-phase detection system to identify pathological gait. *IEEE Trans. Inf. Technol. Biomed.* 14, 1173–1179. doi: 10.1109/TITB.2010.2058813
- Shah, V. V., McNames, J., Mancini, M., Carlson-Kuhta, P., Spain, R. I., Nutt, J. G., et al. (2020). Laboratory versus daily life gait characteristics in patients with multiple sclerosis, Parkinson's disease, and matched controls. *J. Neuroeng. Rehabil.* 17, 1–12. doi: 10.1186/s12984-020-00781-4
- Stolze, H., Petersen, G., Raethjen, J., Wenzelburger, R., and Deuschl, G. (2001). The gait disorder of advanced essential tremor. *Brain* 124, 2278–2286. doi: 10.1093/brain/124.11.2278
- Tiwari, A., and Joshi, D. (2020). An infrared sensor-based instrumented shoe for gait events detection on different terrains and transitions. *IEEE Sens. J.* 20, 10779–10791. doi: 10.1109/JSEN.2020.2994956
- Tunca, C., Pehlivan, N., Ak, N., Arnrich, B., Salur, G., and Ersoy, C. (2017). Inertial sensor-based robust gait analysis in non-hospital settings for neurological disorders. *Sensors* 17, 825. doi: 10.3390/s17040825
- Ullrich, M., Küderle, A., Hannink, J., Del Din, S., Gaßner, H., Marxreiter, F., et al. (2020). Detection of gait from continuous inertial sensor data using harmonic frequencies. *IEEE J. Biomed. Health Inform.* 24, 1869–1878. doi: 10.1109/JBHI.2020.2975361
- Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., et al. (2020). Scipy 1.0: fundamental algorithms for scientific computing in python. *Nat. Methods* 17, 261–272. doi: 10.1038/s41592-019-0686-2
- Vu, H. T. T., Dong, D., Cao, H.-L., Verstraten, T., Lefeber, D., Vanderborght, B., et al. (2020). A review of gait phase detection algorithms for lower limb prostheses. *Sensors* 20, 3972. doi: 10.3390/s20143972
- Warmerdam, E., Hausdorff, J. M., Atrsaei, A., Zhou, Y., Mirelman, A., Aminian, K., et al. (2020). Long-term unsupervised mobility assessment in movement disorders. *Lancet Neurol.* 19, 462–470. doi: 10.1016/S1474-4422(19)30397-7
- Wu, J., Kuruvithadam, K., Schaefer, A., Stoneham, R., Chatzipirpiridis, G., Easthope, C. A., et al. (2021). An intelligent in-shoe system for gait monitoring and analysis with optimized sampling and real-time visualization capabilities. *Sensors* 21, 2869. doi: 10.3390/s21082869
- Yu, H., Wang, D.-H., Yang, C.-J., and Lee, K.-M. (2010). "A walking monitoring shoe system for simultaneous plantar-force measurement and gait-phase detection," in *2010 IEEE/ASME International Conference on Advanced Intelligent Mechatronics* (Montreal, QC: IEEE), 207–212.
- Zahradka, N., Verma, K., Behboodi, A., Bodt, B., Wright, H., and Lee, S. C. (2020). An evaluation of three kinematic methods for gait event detection compared to the kinetic-based 'gold standard'. *Sensors* 20, 5272. doi: 10.3390/s20185272

Zhang, H., Guo, Y., and Zanotto, D. (2019). Accurate ambulatory gait analysis in walking and running using machine learning models. *IEEE Trans. Neural Syst. Rehabil. Eng.* 28, 191–202. doi: 10.1109/TNSRE.2019.2958679

Zhao, H., Wang, Z., Qiu, S., Shen, Y., and Wang, J. (2017). “IMU-based gait analysis for rehabilitation assessment of patients with gait disorders,” in *2017 4th International Conference on Systems and Informatics (ICSAI)* (Hangzhou: IEEE), 622–626.

Zhen, T., Yan, L., and Yuan, P. (2019). Walking gait phase detection based on acceleration signals using lstm-dnn algorithm. *Algorithms* 12, 253. doi: 10.3390/a12120253

Zhu, S., Anderson, H., and Wang, Y. (2012). “A real-time on-chip algorithm for imu-based gait measurement,” in *Pacific-Rim Conference on Multimedia* (Springer: Singapore), 93–104.